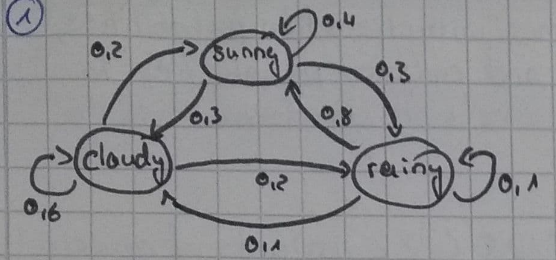
structural bioinformatics - assignment 03

01.12.21

Talha Rehman

Johanna Becher

①



transition probability matrix:

$$\begin{array}{c c} & \begin{array}{ccc} \text{sunny} & \text{cloudy} & \text{rainy} \end{array} \\ \begin{array}{c} \text{sunny} \\ \text{cloudy} \\ \text{rainy} \end{array} & \begin{pmatrix} 0{,}4 & 0{,}2 & 0{,}8 \\ 0{,}3 & 0{,}6 & 0{,}1 \\ 0{,}3 & 0{,}2 & 0{,}1 \end{pmatrix} = A \end{array}$$

We discussed a lot about this exercise and found two ways to interprete it:

1) We want to know the probability, that the weather on the 5th day from today is sunny, given that the other 4 days followed the given path and that today is sunny.

today = sunny path = rainy, rainy, cloudy, sunny, sunny

Since we are looking at a markov model, the weather at day i only depends on the weather at day i-1.
we can calculate the probability as follows

$$p = p(\text{rainy}\,|\,\text{sunny}) * p(\text{rainy}\,|\,\text{rainy}) * p(\text{cloudy}\,|\,\text{rainy}) * p(\text{sunny}\,|\,\text{cloudy}) * p(\text{sunny}\,|\,\text{sunny})$$
$$= 0{,}3 * 0{,}1 * 0{,}1 * 0{,}2 * 0{,}4$$
$$= 0{,}00024$$

2) We want to know the probability that the weather on day x is y, only depending on the fact, that today is sunny.

$X_s^{(k)}$ = probability, that it will be sunny k days from now on, no matter what happened in the time in between

$$x^{(k+1)} = A\,x^{(k)}, \quad x^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

day	0	1	2	3	4	5
weather	s	r	r	c	s	s
	$\begin{pmatrix} \boxed{1} \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0{,}4 \\ 0{,}3 \\ \boxed{0{,}3} \end{pmatrix}$	$\begin{pmatrix} 0{,}46 \\ 0{,}33 \\ \boxed{0{,}21} \end{pmatrix}$	$\begin{pmatrix} 0{,}418 \\ \boxed{0{,}357} \\ 0{,}225 \end{pmatrix}$	$\begin{pmatrix} \boxed{0{,}4186} \\ 0{,}3621 \\ 0{,}2193 \end{pmatrix}$	$\begin{pmatrix} \boxed{0{,}4153} \\ 0{,}36477 \\ 0{,}21993 \end{pmatrix}$

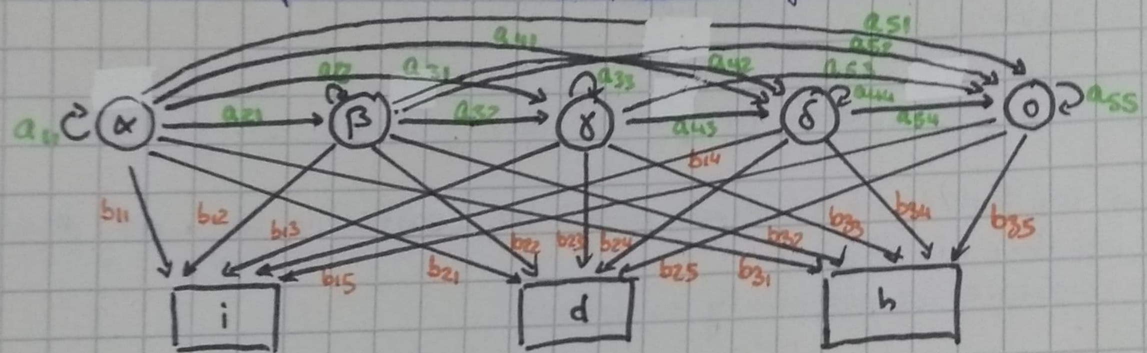
In □ are the probabilities for the different days and the weather on that day labelled.
The probability, that the weather is rainy on the day after tomorrow, given that today is sunny is 0,21.

②

A hidden markov model is a markov process with observable and hidden components.
Every hidden component or state comes with some transmission probabilities to change into an other hidden state and some emission probabilities, to produce an observable state.

In our example, we thought of using the SARS-CoV 2 variants of concern (as stated by WHO) as hidden states and different rates as infected, deaths or hospitalized as observable states.
To make things a bit easier we assume that mutations can't be created backwards (β can't mutate back to α).

With those informations our model looks as follows:



α = Alpha, B.1.1.7
β = Beta, B.1.351
γ = Gamma, P.1
δ = Delta, B.1.617.2
O = Omikron, B.1.1.529

i = infected / 100.000 residents / 7 days
d = deaths / "
h = hospitalized / "

5 hidden states → transition prob. matrix

$$\begin{array}{c|ccccc} & \alpha & \beta & \gamma & \delta & O \\ \hline \alpha & a_{11} & 0 & 0 & 0 & 0 \\ \beta & a_{21} & a_{22} & 0 & 0 & 0 \\ \gamma & a_{31} & a_{32} & a_{33} & 0 & 0 \\ \delta & a_{41} & a_{42} & a_{43} & a_{44} & 0 \\ O & a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{array} = A$$

3 observable states → emission prob. matrix

$$\begin{array}{c|ccccc} & \alpha & \beta & \gamma & \delta & O \\ \hline i & b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ d & b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ h & b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \end{array} = B$$